

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-039432

(43)Date of publication of application : 12.02.1999

(51)Int.Cl.

G06K 9/68

(21)Application number : 09-190801

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 16.07.1997

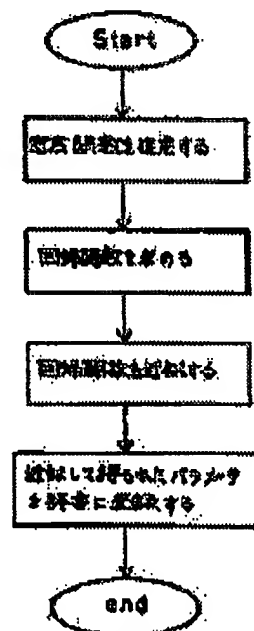
(72)Inventor : MIZUTANI HIROYUKI

(54) METHOD AND DEVICE FOR CHARACTER RECOGNITION

(57)Abstract:

PROBLEM TO BE SOLVED: To reduce an operation quantity regarding recognition by setting the discrimination border of a category according to a distribution of recursive function values and discriminating the category of an inputted character pattern on the basis of the set discrimination border.

SOLUTION: Probability density functions are estimated by the categories of sets of samples of previously given character patterns and used to obtain recursive function composed of representative vector values indicating the respective categories that the character patterns belong to. Then the parameters of the obtained recursive function are set to approximated recursive function values, which are registered in a dictionary. Further, the discrimination border of the categories is set according to the distribution of the registered recursive function values and on the basis of the set discrimination border, the categories of the inputted character patterns are discriminated.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-39432

(43) 公開日 平成11年(1999) 2月12日

(51) Int.Cl.^a

G 0 6 K 9/68

識別記号

F I

G 0 6 K 9/68

E

審査請求 未請求 請求項の数 2 O L (全 6 頁)

(21) 出願番号 特願平9-190801

(22) 出願日 平成9年(1997) 7月16日

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 水谷 博之

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

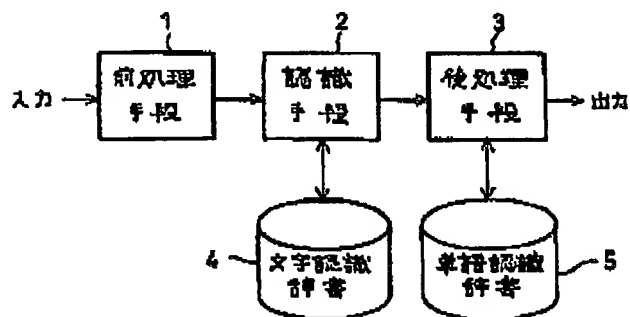
(74) 代理人 弁護士 外川 英明

(54) 【発明の名称】 文字認識方法及び装置

(57) 【要約】

【課題】本発明は、従来の多峰性の分布形状をもつ複雑な密度関数を用いた場合に伴う演算量の増加を抑えるとともに、次元圧縮の際にパターン分布の高次元性の為に生じる適切なモデル当てはめの困難さを解消し、学習サンプルに対してクラスタ解析等のアドホックな処理を行わなくても、分布に応じた識別境界を自動的に決定することを目的とする。

【解決手段】本発明は、予め与えられた文字パターンのサンプルの集合から各カテゴリごとに確率密度関数を推定し、推定された前記確率密度関数を用い、前記文字パターンの属する各カテゴリを示す代表ベクトル値により構成された回帰関数を求め、求められた前記回帰関数のパラメータを近似した回帰関数値を求め、求められた回帰関数値の分布に応じてカテゴリの識別境界を設定し、設定された前記識別境界をもとに、入力された文字パターンのカテゴリの識別を行なうことを特徴とする。



【特許請求の範囲】

【請求項 1】 予め与えられた文字パターンのサンプルの集合の各カテゴリごとに確率密度関数を推定し、推定された前記確率密度関数を用い、前記文字パターンの属する各カテゴリを示す代表ベクトル値により構成された回帰関数を求め、

求められた前記回帰関数のパラメータを近似した回帰関数値を求め、

求められた回帰関数値の分布に応じてカテゴリの識別境界を設定し、

設定された前記識別境界をもとに、入力された文字パターンのカテゴリの識別を行なうことを特徴とする文字認識方法。

【請求項 2】 予め与えられた文字パターンのサンプルの集合の各カテゴリごとに確率密度関数を推定する密度関数推定手段と、

この密度関数推定手段によって推定された前記確率密度関数を用い、前記文字パターンの属する各カテゴリを示す代表ベクトル値により構成された回帰関数を求める回帰関数算出手段と、

この回帰関数算出手段によって求められた前記回帰関数のパラメータを近似した回帰関数値を求める回帰関数値算出手段と、

この回帰関数値算出手段によって求められた回帰関数値の分布に応じてカテゴリの識別境界を設定する設定手段と、

この設定手段によって設定された前記識別境界をもとに、入力された文字パターンのカテゴリの識別を行なう識別手段とを有することを特徴とする文字認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、文字パターンの分布に応じた識別境界を自動的に決定する文字認識方法及び装置に関する。

【0002】

【従来の技術】 文献「パターン認識：森健一監修、電子情報通信学会、1993年」によると、ペイズ識別で代表される統計的決定理論を用いた文字認識では、原文字パターンから特徴抽出を経て変換された特徴パターンが構成するパターン空間において、パターンベクトルの確率密度関数（分布密度）が知られれば未知文字パターンの分類を行なうことが出来る。従来の統計的決定理論では、確率密度関数として正規分布を仮定することが多かった。

【0003】 しかしながら、この確率密度関数は特徴抽出に応じて変化すると考えるのが自然であり、必ずしも正規分布が適切な分布であるとは考えにくい。この視点にたち、多峰性の分布を学習サンプルから推定してパターン認識を行なう手法が提案されているが、この手法では一般にパラメータ数が多くなり、認識率の割には認識

に関わる演算量が多くなり、実用的なシステムには不向きであるという問題があった。

【0004】 また、部分空間法で代表されるように、学習サンプルに対して主成分分析により次元圧縮を行う手法は、パターン空間におけるパターンベクトルの分布に対して超平面を当てはめるものであるが、一般には分布の形状に応じて適切な当てはめモデルを導入した方がより次元圧縮の効果が期待される。

【0005】 従って当てはめモデルを導入して適切な次元圧縮を行おうとするには、用いる特徴に応じて事前にパターン空間におけるパターンベクトルの分布を調べる必要がある。しかしながらパターン空間は一般に高次元空間であるため、分布の形状を調査することは大変難しいという問題があった。

【0006】 一方、パターン空間において学習サンプルからカテゴリの境界を表す識別関数を求め、これを元に未知パターンの分類を行なう手法がある。この手法では、識別関数として線形関数あるいは区分的線形関数や非線形関数を用いているが、これら識別関数の決定はあらかじめ学習サンプルをクラスタ解析することにより、予め幾つかの部分群に分けるなどの操作を必要とすることが多く、関数を決定することに対してアドホックなアルゴリズムを導入することが多く、処理が煩雑になるとともに、得られた識別関数の最適性について議論することは難しかった。

【0007】

【発明が解決しようとする課題】 そこで本発明は、従来の多峰性の分布形状をもつ複雑な密度関数を用いた場合に伴う演算量の増加を抑えるとともに、次元圧縮の際パターン分布の高次元性の為に生じる適切なモデル当てはめの困難さを解消し、学習サンプルに対してクラスタ解析等の処理を行わなくても、分布に応じた識別境界を自動的に決定することを目的とする。

【0008】

【課題を解決するための手段】 本発明は、予め与えられた文字パターンのサンプルの集合の各カテゴリごとに確率密度関数を推定し、推定された前記確率密度関数を用い、前記文字パターンの属する各カテゴリを示す代表ベクトル値により構成された回帰関数を求め、求められた前記回帰関数のパラメータを近似した回帰関数値を求め、求められた回帰関数値の分布に応じてカテゴリの識別境界を設定し、設定された前記識別境界をもとに、入力された文字パターンのカテゴリの識別を行なうことを特徴とするものである。

【0009】

【発明の実施の形態】 本発明は、離散的に与えられた文字のサンプル集合からカテゴリ毎に少なくとも1つの連続な確率密度関数を推定し、従来の回帰分析で必要だった当てはめモデルを用いることなく、推定された密度関数を用いるだけで、各カテゴリを指し示すよう付加され

た代表ベクトル値を回帰する回帰関数を求める。

【0010】また、この回帰関数を少数パラメータで近似することにより、認識処理に係わる演算量を軽減させるとともに、クラスタ解析等の処理を行うこと無しに、学習サンプルに対する回帰関数値と代表ベクトルとの2乗誤差を最小化させる自動学習により分布の形状に応じた識別境界を構成する。

【0011】そして認識は、近似された回帰関数の未知パターンに対するベクトル関数値を計算し、各成分値の最大値をとる成分と1対1に対応したカテゴリへ未知パターンの類別を行なうものである。以下、図面をもとに本発明を説明する。

【0012】図1は、文字認識装置の全体構成図である。前処理手段1は、スキャナやタブレット等の入力手段から入力された文字データの文字区切りやノイズ消去、文字の大きさの正規化等を行う。

【0013】認識手段2は、前処理手段1において処理された文字列を、文字認識辞書4をもとにマッチング処理を行い認識する。次に、後処理手段3は、認識手段2において認識された文字列を、単語認識辞書5をもとに認識単語を導き出し出力する。ここで、文字認識辞書4と単語認識辞書5を1つの辞書として構成しても良い。

【0014】本発明は、認識手段2における認識処理の一部であり、図2に示したように4段階の手順からなる。また認識に用いられる認識辞書は図3に示したように4段階の手順を経て作成される。はじめに認識辞書作成手順を説明する。

【0015】確率変数 $\{x, y\}$ と確率密度 $\rho(x, y)$ が与えられた時、 y を回帰する x に対する回帰関数 $f(x)$ は一般回帰によると次式で求まる。(ここで、 x, y, f はベクトル値である。)

【0016】

【数1】

$$\bar{f}(\bar{x}) = \frac{\int_{-\infty}^{\infty} \bar{y} \rho(\bar{x}, \bar{y}) d\bar{y}}{\int_{-\infty}^{\infty} \rho(\bar{x}, \bar{y}) d\bar{y}}$$

【0017】今、サンプルパターン数を N 、このうち i 番目のパターンを x_i (ベクトル値) とし、このパターンの属するカテゴリを指標する指標ベクトルを y_i と書く。以下では、密度関数 ρ の具体例として、多次元正規分布から構成される混合分布を用いるが、一般的には

【数1】の右辺の各積分が存在し、かつ右辺が有限となる如何なる密度関数 ρ を用いてもよい。混合分布を用いた場合、密度関数 ρ は

【0018】

【数2】

$$\rho(\bar{x}, \bar{y}) = \frac{1}{N} \sum_{i=1}^N G(\bar{x} - \bar{x}_i, \Sigma_i) G(\bar{y} - \bar{y}_i, \Sigma_i)$$

である。ここで $G(x, \Sigma)$ は平均が0、共分散行列を

Σ とした多次元正規分布である。【数1】に【数2】を代入すると、具体的に近似した回帰関数として

【0019】

【数3】

$$\bar{f}(\bar{x}) = \frac{\sum_{i=1}^N G(\bar{x} - \bar{x}_i, \Sigma_i) \bar{y}_i}{\sum_{i=1}^N G(\bar{x} - \bar{x}_i, \Sigma_i)}$$

が得られる。

【0020】次に、回帰関数 f を m 組のパラメータ $\{\alpha_i, \beta_i, \Sigma_i\}$ (i は $1, \dots, m$) を用いて近似する(ここで、 α, β はベクトル値)。これら m 組のパラメータ $\{\alpha_i, \beta_i, \Sigma_i\}$ のうち、 α を第1のパラメータ、 β を第2のパラメータ、 Σ を第3のパラメータとする。

【0021】本実施例では、【数3】で表された f を直接関数近似するのではなく、サンプルデータと近似された回帰関数 f^* との2乗誤差を最小にするパラメータの組を求めることで近似を行なう。また、 m 組のパラメータ $\{\alpha_i, \beta_i, \Sigma_i\}$ の内、簡単のために $\Sigma_1 = \Sigma_2 = \dots = \Sigma_m = \Sigma$ とし、 Σ の値は予め設定されているものとする。この場合、近似される回帰関数 f^* は m 組のパラメータ $\{\alpha_i, \beta_i\}$ (i は $1, \dots, m$) を有する。この場合、サンプルデータと近似関数との2乗誤差 E は、

【0022】

【数4】

$$E(\{\bar{\alpha}_i, \bar{\beta}_i\}_{i=1}^m) = \sum_{i=1}^N \|\bar{y}_i - \bar{f}^*(\bar{x}_i)\|^2$$

となる。そして、変分法により $E(\{\alpha_i, \beta_i\})$ を最小にするパラメータの組は次式の連立方程式を満足する。

【0023】

【数5】

$$\sum_{i=1}^N \frac{\partial \bar{f}^*(x_i)}{\partial \bar{\alpha}_k} (\bar{y}_i - \bar{f}^*(\bar{x}_i)) = 0$$

$$\sum_{i=1}^N \frac{\partial \bar{f}^*(x_i)}{\partial \bar{\beta}_k} (\bar{y}_i - \bar{f}^*(\bar{x}_i)) = 0$$

【0024】これら m 組の連立方程式を各パラメータに関して解けば2乗誤差を最小にする意味で最適なパラメータが求まる。しかしこれらの連立方程式は、一般には非線形連立方程式となり解析解を求めることは難しい。

【0025】また、各パラメータの次元が大きい場合、数値的に連立方程式を解くことも困難である。ここでは一例として最急降下法による数値解法を示す。最急降下法では学習の進行の具合を調整する任意の正実数 ε を導入して、次式のようにパラメータを逐次更新する。

【0026】

【数6】

$$\begin{aligned}\alpha_k^{(i+1)} &= \alpha_k^{(i)} - \frac{c}{2} \frac{\partial E}{\partial \alpha_k^{(i)}} \\ &= \alpha_k^{(i)} - c \sum_{i=1}^N \frac{\partial \tilde{f}^*(x_i)}{\partial \alpha_k^{(i)}} (\tilde{y}_i - \tilde{f}^*(\tilde{x}_i))\end{aligned}$$

【0027】

【数7】

$$\alpha_k^{(i+1)} = \alpha_k^{(i)} + c \Sigma^{-1} \sum_{i=1}^N \frac{G_{ik}}{M_{\alpha i}} \langle \tilde{\beta}_k^{(i)} - \tilde{f}_i^{(i)} | \tilde{y}_i - \tilde{f}_i^{(i)} \rangle (\tilde{x}_i - \alpha_k^{(i)})$$

【0029】

【数8】

$$\tilde{\beta}_k^{(i+1)} = \tilde{\beta}_k^{(i)} + c \sum_{i=1}^N \frac{G_{ik}}{M_{\alpha i}} (\tilde{y}_i - \tilde{f}_i^{(i)})$$

$$\tilde{f}_i = \frac{\sum_{k=1}^n \tilde{\beta}_k G_{ik}}{\sum_{k=1}^n G_{ik}}, \quad M_{\alpha i} = \sum_{k=1}^n G_{ik}, \quad G_{ik} = G(\tilde{x}_i - \alpha_k, \Sigma),$$

および、 $\langle \rangle$ はベクトルの内積を表す。

【0031】学習はパラメータの初期値を適切に設定した後、上の更新規則で各パラメータを更新し、【数4】で計算されるEの減少率がある一定の閾値を割った時点終了させればよい。

【0032】学習時のパラメータ初期値設定に関しては、最適化問題が非線形最適化の場合は最適化アルゴリズムによっては学習の効率化や解の最適性に影響することが知られており重要である。本実施例では一例として、サンプルデータの重心を出発点とした代表パターン選出する手順を示す。この処理は各カテゴリごとに独立に行なう。

【0033】まず、サンプルパターンから重心を求める。次に、サンプルパターンから各重心から最も近い代表パターンを選出する。次に、各サンプルパターンと代表パターンのユークリッド距離を計算し、ユークリッド距離を合計し記憶する。次に、合計されたユークリッド距離が最大となるサンプルパターンを代表パターンとする。そして、必要な代表パターン数が得られれば、これらを $\{\alpha i\}$ の初期値として登録し終了する。そうでなければ3に戻る。

【0034】以上の手法以外に、例えばファジークラスタリングやk-meanクラスタリング等の手法によりサンプルデータから各カテゴリを代表する代表パターンを選出し、 $\{\alpha i\}$ の初期値としてもよい。

【0035】認識用辞書には、 $\{\alpha i, \beta i, \Sigma i\}$ を登録する。本実施例では Σ を共分散行列としたが、パラメータ数を減らす目的で、例えば、単位行列 I および正実数 σ を用いて、

【0036】

$$\begin{aligned}\tilde{\beta}_k^{(i+1)} &= \tilde{\beta}_k^{(i)} - \frac{c}{2} \frac{\partial E}{\partial \tilde{\beta}_k^{(i)}} \\ &= \tilde{\beta}_k^{(i)} - c \sum_{i=1}^N \frac{\partial \tilde{f}^*(x_i)}{\partial \tilde{\beta}_k^{(i)}} (\tilde{y}_i - \tilde{f}^*(\tilde{x}_i))\end{aligned}$$

【数6】と【数7】で f^* を具体的に書き下すと、学習アルゴリズムは次のようになる。

【0028】

【数8】

である。ここで、

【0030】

【数10】

【数11】

$$\Sigma = \sigma^2 I$$

と表される Σ を用いることも出来る。また、指標ベクトルの各成分とカテゴリを1対1に対応させ、対応した成分の値を1、それ以外の成分を0と設定すると、 $\{\beta i\}$ を上記学習アルゴリズムで学習しなくてもよい。

【0037】次に認識手順について説明する。まず、認識辞書から $\{\alpha i, \beta i, \Sigma i\}$ を読み込み、回帰関数を【数3】の形式で再構成する。すなわち、未知のパターンを x として、

【0038】

【数12】

$$\tilde{f}(x) = \frac{\sum_{i=1}^m G(x - \alpha_i, \Sigma_i) \tilde{\beta}_i}{\sum_{i=1}^m G(x - \alpha_i, \Sigma_i)}$$

となる。

【0039】今、カテゴリ数を n としたとき、未知のパターンを x に対する回帰関数 $f(x)$ のうち、 i 番目のカテゴリ $c i$ に対する成分を $f_{c i}$ と書く。このときカテゴリへ類別は $C(\cdot)$ を類別演算として

【0040】

【数13】

$$C(x) = c_i \quad \text{if} \quad f_{c_i} = \max_j \{f_{c_j}\}$$

である。

【0041】そしてこの場合、回帰関数の各成分の値は未知パターン x が各カテゴリに属する確率を表すことを次に示す。ここで、 n 個のカテゴリのうち、 i 番目のカテゴリ $c i$ に対する確率密度関数を $p_{c i}$ と書くと、回帰関数 f の各成分 $f_{c i}$ は次式で表現される。

【0042】

【数14】

$$f_{ci}(\vec{x}) = \frac{\rho_{ci}(\vec{x})}{\sum_{i=1}^n \rho_{ci}(\vec{x})}$$

【0043】この式から $f_{ci}(\vec{x})$ の値は、ある \vec{x} における全カテゴリの確率密度に対する c_i カテゴリの確率密度の比を表している。従ってこの密度比の大きさは、 \vec{x} における識別結果の信頼度（統計的な確からしさ）を表している。従って、この値を用いて識別結果を棄却するリジェクト処理が容易に行える。

【0044】リジェクト処理としては、例えばある一定の閾値を設け、この閾値よりも信頼度が低ければ識別結果を棄却すればよい。図4は類別すべきカテゴリが3つある場合の識別関数を説明する概念図である。この図ではパターン空間は2次元であり、3つの大きな黒点は3つのパラメータ α を表している。また、 f_A 、 f_B 、 f_C は回帰関数の各成分を表している。すなわち、

【0045】

【数15】

$$\vec{f}(\vec{x}) = (f_A(\vec{x}), f_B(\vec{x}), f_C(\vec{x}))$$

である。この時、識別境界は

【0046】

【数16】

$$f_A(\vec{x}) = f_B(\vec{x}) = f_C(\vec{x})$$

を満足する。図中、太い実線で識別境界を示した。細い実線および破線は等高線を表す。

【0047】さて、【数14】を用いて【数16】に対応する方程式を導き、確率密度関数として等方的で均一な正規分布を仮定すると、識別境界上の関数値は3直線の交点では $1/3$ であり、交点から離れるにしたがって次第に $1/2$ に近付くことが示される。

【0048】また、境界から離れた位置では距離が遠ざかるに従い各成分の値は 1 に漸近することが示される。この例の場合リジェクト処理としては、例えば回帰関数値の最大成分値が 0.6 未満の場合は識別結果を棄却すればよい。

【0049】別なりリジェクト処理例として、回帰関数値の最大成分値だけではなく、最大成分値を含む複数の成

分値を用いることも出来る。例えば最大成分値と第二最大成分値とからなる2次元ベクトルを構成し、これに対して予め2次元の棄却領域を設け、この2次元ベクトルが棄却領域内の値を取った場合は棄却するといった処理を行なうことが考えられる。

【0050】また、本願発明の別の実施形態として、予め与えられた文字パターンのサンプルの集合の各カテゴリごとに確率密度関数を推定し、推定された前記確率密度関数を用い、前記文字パターンの属する各カテゴリを示す代表ベクトル値により構成された回帰関数を求め、求められた前記回帰関数のパラメータを近似した回帰関数値を求め、求められた回帰関数値の分布に応じてカテゴリの識別境界を設定し、設定された前記識別境界をもとに、入力された文字パターンのカテゴリの識別を行なう手順で実現されるプログラムをコンピュータで読み込み可能なフロッピーディスクやCD-ROMなどの記録媒体として実施することも可能である。

【0051】

【発明の効果】以上説明したように本発明によれば、当てはめモデルを用いることなく学習サンプルから推定された密度関数に基づいて識別関数を構成し、この識別関数を適切なパラメータ数で近似することにより、認識に係わる演算量を軽減させると共に分布に応じた識別境界を自動学習により構成することが出来るものである。さらに、識別結果に対する統計的な信頼度を表すように回帰関数を構成することが容易であり、これに伴い識別結果を棄却するリジェクト処理が容易になる。

【図面の簡単な説明】

【図1】本願発明の一実施形態に関わる文字認識装置の全体構成を示す図である。

【図2】本願発明の一実施形態に関わる文字認識装置の認識手順を示す図である。

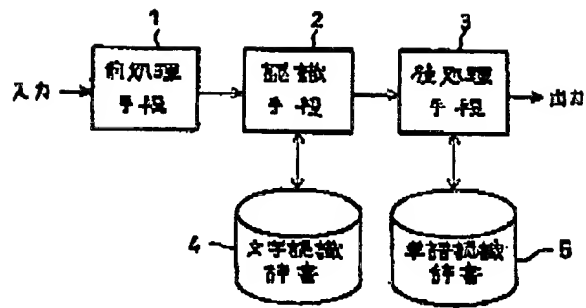
【図3】本願発明の一実施形態に関わる文字認識装置の認識辞書を作成する手順を示す図である。

【図4】本願発明の一実施形態に関わる類別すべきカテゴリが3つある場合の識別関数を説明する図である。

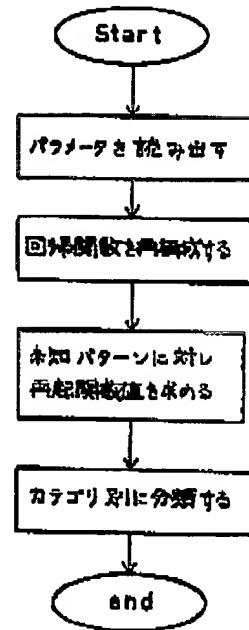
【符号の説明】

- 1 前処理手段、 2 認識手段、 3 後処理手段、
- 4 文字認識辞書
- 5 単語認識辞書

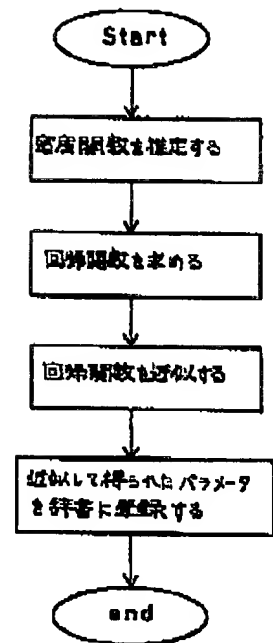
【図 1】



【図 2】



【図 3】



【図 4】

